Learning an Optimized Deep Neural Network for Link Prediction on Knowledge Graphs

Wilcke, W.X.

2015

Learning an Optimized Deep Neural Network for Link Prediction on Knowledge Graphs

Xander Wilcke
w.x.wilcke@vu.nl
VU University Amsterdam, the Netherlands

VU UNIVERSITY AMSTERDAM

Aim of this Research

Recent years have seen the emergence of graph-based Knowledge Bases build upon **Semantic Web** technologies, known as **Knowledge Graphs**. Effectively learning from these complex relational structures remains a challenge yet to be overcome.

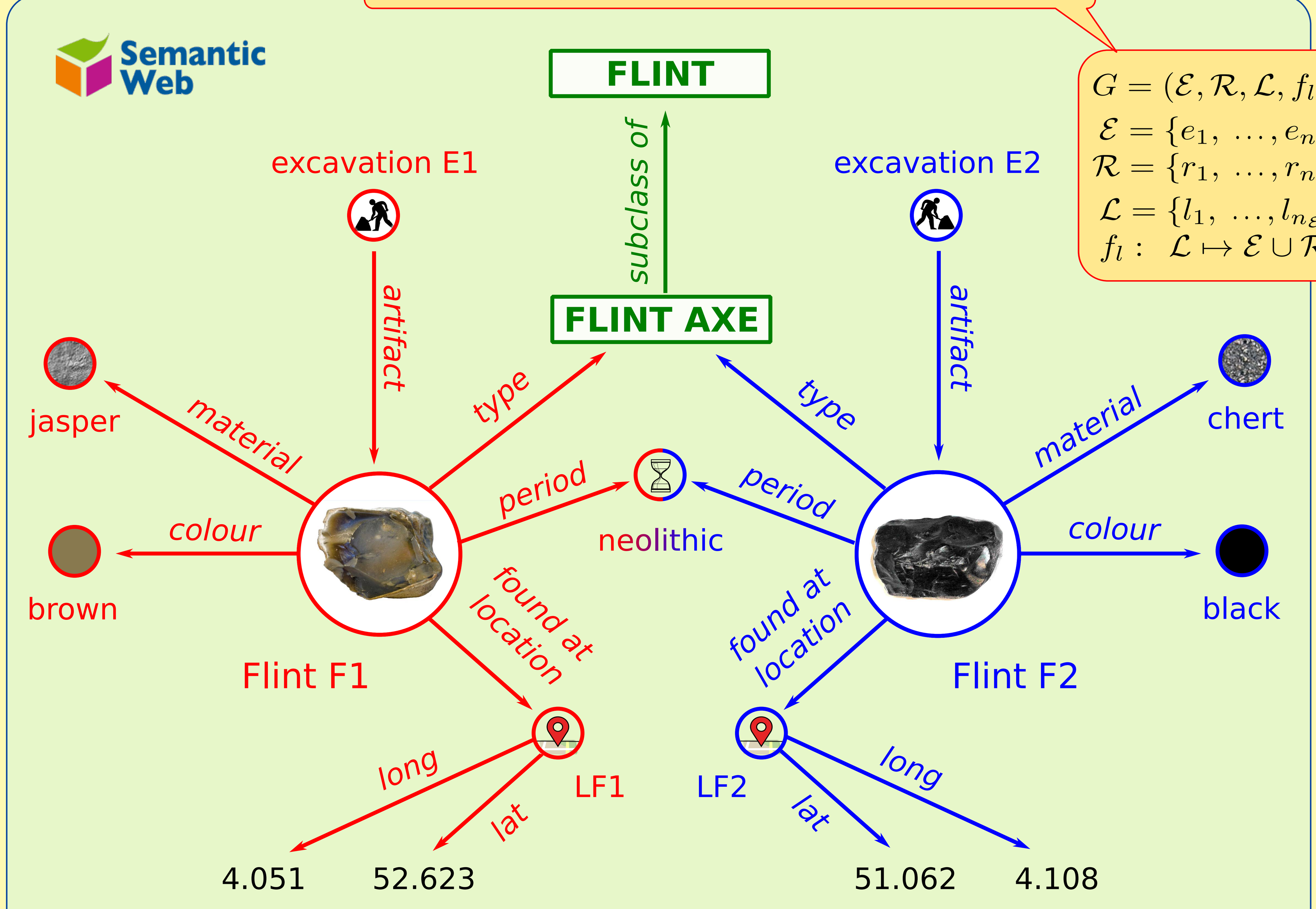
For this purpose, we are investigating the effectiveness of **Link Prediction** through means of **Deep Learning an Artificial Neural Network**, as well to optimize both learning method and model through **Bayesian Hyper-parameter optimization**. Moreover, during evaluation, special attention will be given to the usefulness of made predictions to domain experts.

Knowledge Graphs

- Describe **factual information as relations between entities** (edges between vertices)
- Relations and entities are assigned **special labels which guard their semantics**
- Semantic labels are strictly defined in common ontologies (schemes)
- **Ambiguity is minimized** within and between Knowledge Graphs by freely sharing ontologies on the (Semantic) Web
- Inherit **Deductive Reasoning capabilities** from their underlying formal system

Example

A Knowledge Graph contains a finite set of entities and a finite set of relations, as well as a finite set of labels and a corresponding mapping over the graph's elements.



$G = (\mathcal{E}, \mathcal{R}, \mathcal{L}, f_l)$
$\mathcal{E} = \{e_1, \ldots, e_{n_\mathcal{E}}\}$
$\mathcal{R} = \{r_1, \ldots, r_{n_\mathcal{R}}\}$
$\mathcal{L} = \{l_1, \ldots, l_{n_{\mathcal{E}\cup\mathcal{R}}}\}$
$f_l : \mathcal{L} \mapsto \mathcal{E} \cup \mathcal{R}$

A small Knowledge Graph depicting two distinct flints (red and blue) found during two separate archaeological excavations. Both flints were found to be of the *flint axe* class (rectangles), which is a subclass of the *flint* class. Each flint has various properties (cursive labels), which may take a value from a finite (circles) or infinite set (black text). Some relations might be left undefined, as their existence is simply unknown due to the **Open World Assumption**.

Queries fired over these data are able to answer questions like :

- List all flints of type *flint axe* which are of colour black and from period neolithic.
- List all brown flints found within 10 m from a black flint from the same period.
- List all excavations in the Netherlands during which chert *flint axes* were found.

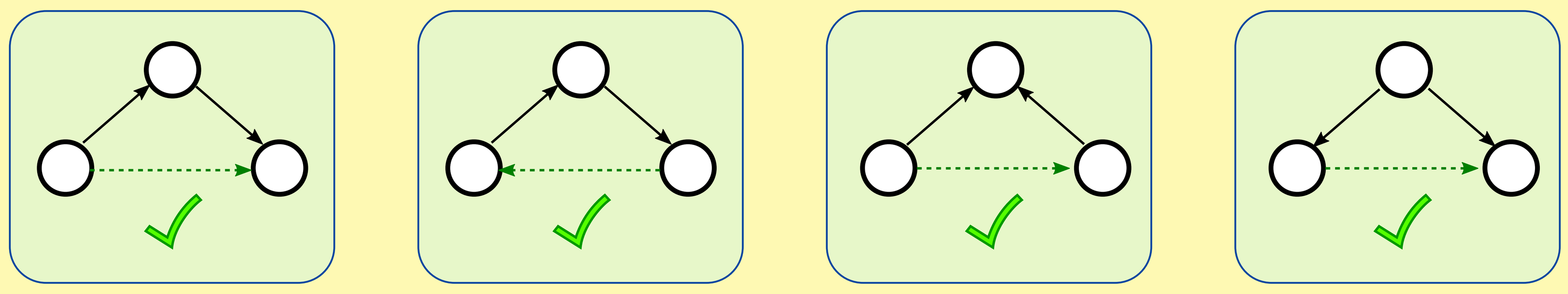
Motivation

By default, reasoning over Knowledge Graphs is

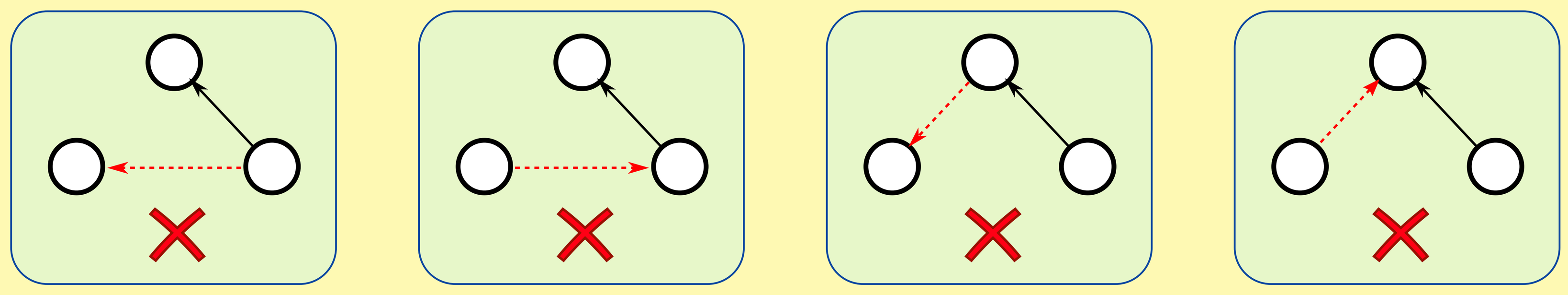
- **completely dependent on axiomatic prior knowledge, and**
- **solely able to derive information that was already implicitly present.**

Hence, the existence of many relations (= facts) remains unknown

Inferable relations :



Non-inferable relations :



Using Link Prediction, we can estimate these relations' probability of existence

Knowing which relations are likely to exist is **highly welcomed by Domain Experts** :

- as starting points from which they form new potential hypotheses
- as potential evidence to support proving of disproving existing hypotheses
- as measure to evaluate the trustworthiness of potential sources of information

Methodology

Propositionalization Strategy

Generated features hold information on :

- the **labels** of the sample's elements (embedded features)
- the **semantics** of the sample's elements (ontological features)
- the **Local Neighbourhood** up to depth *d* of the sample's entities (graph features)

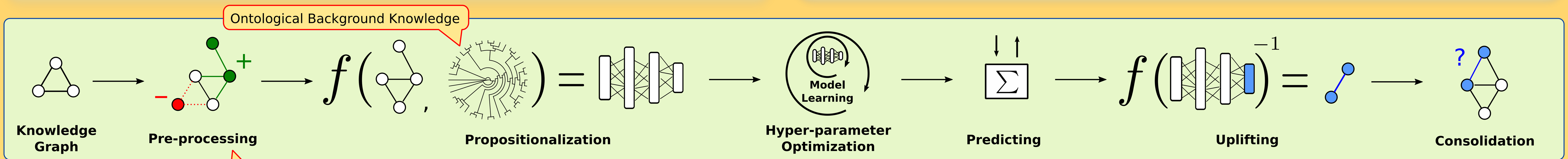
Learning and Optimization

Training the Neural Network involves :

- **pre-training** the weight matrices layer-wise using **Restricted Boltzmann machines**
- **finetuning** the weight matrices using **supervised back-propagation** during online learning
- **Bayesian Hyper-parameter Optimization** using **Random Forests** as prior

Evaluating Predictions

- area under the precision-recall curve (**AUC-PR**)
- **usefulness of predictions to domain experts** (qualitative measures)



Rationale

We have chosen a state-of-the-art **feedforward Neural Network** for our model.
We motivate this decision as follows :

- Being a **latent-feature model**, they are a sensible choice to learn from real-world data, due to their **robustness** towards noise and inconsistencies, as well their ability to cope with large-scale and high-dimensional data sets.
- Neural Networks are **universal function approximators**, which allows them to model any arbitrary relational structure, given enough model complexity.
- Recent breakthroughs have made it possible to **learn deep architectures**, which radically improves their ability to solve complex learning problems.

Forthcoming Research

Minimizing the information loss cause by the propositionalization process is a continuing effort. At present, this loss mainly involves ontological features. Hence, integration of these features will be our next point of focus. For this purpose, we intent to **investigate the trade-off between generic and domain-specific exploitation of ontological features**.

In the near future, we intent to **extensively evaluate our approach on real-world Knowledge Graphs**, and to relate it to comparable methods in recent literature. To this end, we are currently working together with interested parties from Digital Humanities, who are providing us with data, as well as with qualitative standards to which the usefulness of made predictions will be measured.

Acknowledgements: This research is partially funded by the ARIADNE project through the European Commission under the Community's Seventh Framework Programme, contract no. FP7-INFRASTRUCTURES-2012-1-313193.

NETWORK INSTITUTE
SPIN LAB SPATIAL INFORMATION LABORATORY
Universiteit Leiden
SEVENTH FRAMEWORK PROGRAMME
ARIADNE
www.ariadne-infrastructure.eu